

Polynomial processes and their applications to mathematical finance

Christa Cuchiero · Martin Keller-Ressel · Josef Teichmann

Received: 30 December 2008 / Accepted: 19 March 2012 / Published online: 3 July 2012
© Springer-Verlag 2012

Abstract We introduce a class of Markov processes, called m -polynomial, for which the calculation of (mixed) moments up to order m only requires the computation of matrix exponentials. This class contains affine processes, processes with quadratic diffusion coefficients, as well as Lévy-driven SDEs with affine vector fields. Thus, many popular models such as exponential Lévy models or affine models are covered by this setting. The applications range from statistical GMM estimation procedures to new techniques for option pricing and hedging. For instance, the efficient and easy computation of moments can be used for variance reduction techniques in Monte Carlo methods.

Keywords Markov processes · Diffusions with jumps · Affine processes · Analytic tractability · Pricing · Hedging

Mathematics Subject Classification 60J25 · 91B70

JEL Classification C02 · G12

C. Cuchiero

University of Vienna, Faculty of Mathematics, Nordbergstrasse 15, 1090 Wien, Austria
e-mail: christa.cuchiero@univie.ac.at

M. Keller-Ressel

TU Berlin, Fakultät II, Institut für Mathematik, Strasse des 17. Juni 136, 10623 Berlin, Germany
e-mail: mkeller@math.TU-Berlin.DE

J. Teichmann (✉)

ETH Zürich, D-MATH, Rämistrasse 101, 8092 Zürich, Switzerland
e-mail: josef.teichmann@math.ethz.ch

1 Introduction

Pricing and hedging of contingent claims are the crucial computations done within every model in mathematical finance. For European-type claims, this amounts to the computation of the expected value of a functional of the (discounted) price process under some martingale measure. (Partial) Hedging portfolios are then constructed via appropriate derivatives, so-called Greeks, of those expected values with respect to model parameters or to the prices. Let us denote the (discounted) price process at time T , a vector in \mathbb{R}^n , by X_T . We can roughly distinguish three cases of complexity for the mentioned computations:

1. The probability distribution of X_T is analytically known.
2. The characteristic function of X_T is analytically known.
3. The local characteristics of X_T are analytically known.

In the first case, a numerical quadrature algorithm is sufficient for the efficient computation of the contingent claim's price $\mathbb{E}[\phi(X_T)]$, where ϕ denotes some payoff function.

In the second case, variants of Plancherel's theorem are applied in order to evaluate the price functional $\mathbb{E}[\phi(X_T)]$, for instance

$$\mathbb{E}[\phi(X_T)] = \int_{\mathbb{R}^n} \widehat{\phi}(u) \mathbb{E}[\exp(i\langle u, X_T \rangle)] du,$$

where $\widehat{\phi}$ denotes the Fourier transform of the function ϕ . Remark that often modifications of the original payoff function are used to make the Fourier methodology applicable. This is numerically efficient, even though its implementation, in particular the complex integration, can take some time (see e.g. [3]). Also there are different levels of what it means to “know analytically” the characteristic function of X_T . In affine models, for example, one might need to solve a high-dimensional Riccati equation for each $u \in \mathbb{R}^n$ to calculate the characteristic function $u \mapsto \mathbb{E}[\exp(i\langle u, X_T \rangle)]$. In this case, “analytic knowledge” involves some precalculations, which also have to be performed efficiently.

The third case is characterized by the use of Monte Carlo methods: one samples from the (unknown) distribution of X_T by generating, for instance through Euler schemes, approximate distributions for X_T . This procedure is very robust, but takes a considerable amount of time. Moreover, the convergence of the Euler scheme requires certain regularity assumptions on the characteristics, e.g. (local) Lipschitz-continuity.

In this article, we should like to add a fourth case which—in the previous order—would correspond to case 1 $\frac{1}{2}$. We can describe a class of Markov processes, called “polynomial processes,” which have the property that the expected value of *any polynomial* of the random variables X_t , $t \geq 0$, is *again a polynomial in the initial value* of the process. This means in particular that moments of all orders of X_T can be computed in an *easy and efficient* way, even though neither its probability distribution nor its characteristic function needs to be known. Loosely speaking, one could say that the expressions for all finite moments are analytically known (up to a matrix exponential).

We shall analyze this class and show that exponential Lévy processes, affine processes or processes of Pearson diffusion type belong to it. The method is best explained by an example. Consider a stochastic volatility model of SVJJ-type [12], i.e., both the logarithmic (discounted) price process and the stochastic volatility can jump. Such models can be described by stochastic differential equations of the type

$$\begin{aligned} dY_t &= (b_1 + \beta_{11} Y_t + \beta_{21} V_t) dt + \sqrt{V_t} dB_{t,1} + dZ_{t,1}, \\ dV_t &= (b_2 + \beta_{22} V_t) dt + \sqrt{V_t} dB_{t,2} + dZ_{t,2}, \end{aligned}$$

where (B_1, B_2) are possibly correlated Brownian motions and (Z_1, Z_2) is a bivariate pure-jump Lévy process, independent of (B_1, B_2) , whose second component Z_2 has positive increments. For such models, there is no easy-to-implement (explicit) formula for the characteristic function, even though they are affine models. Assuming now appropriate moment conditions on the jump measures, the Markov process $X = (Y, V)$ turns out to be a polynomial process, that is, the expected value of any polynomial of (Y_t, V_t) is a polynomial in Y_0 and V_0 . The coefficients of this polynomial can be calculated efficiently by exponentiating a matrix, which can be easily deduced from the (extended) generator. In other words, there is a large subset of claims for which the prices and hedge ratios are explicitly known (up to matrix exponentials). Large can be made precise in the following sense: if the law of X_T , say μ , is characterized by its moments, then “large” means dense, i.e., polynomial claims are dense (with respect to the $L^1(\mu)$ -norm) in the set of “all” claims. If this is not the case, the payoff function can at least be uniformly approximated by polynomials on some interval, which can be chosen according to the support of the probability distribution. The explicit knowledge of prices of polynomial claims then allows to apply variance reduction techniques for Monte Carlo computations.

The remainder of our article is organized as follows. In Sect. 2, we formally introduce the class of m -polynomial processes, establish a relationship to semimartingales and give conditions on the (extended) generator such that a Markov process is m -polynomial. Section 3 deals with examples from the class of m -polynomial processes and Sect. 4 with applications to pricing and hedging in mathematical finance.

2 Polynomial processes

We define polynomial processes as a particular class of time-homogeneous Markov processes with state space $S \subseteq \mathbb{R}^n$, some closed subset of \mathbb{R}^n . To clarify notation, we find it useful to recall the basic ingredients of a time-homogeneous Markov process and the particular assumptions and conventions being made in this article (compare [22, Chap. 3]).

Throughout, S is a closed subset of \mathbb{R}^n and \mathcal{S} denotes its Borel σ -algebra. Since we shall not assume the process to be conservative, we adjoin to the state space S a point $\Delta \notin S$, called cemetery, and set $S_\Delta = S \cup \{\Delta\}$ as well as $\mathcal{S}_\Delta = \sigma(\mathcal{S}, \{\Delta\})$. We make the convention that $f(\Delta) = 0$ for any function f on S .

We consider a time-homogeneous Markov semigroup $(P_t)_{t \geq 0}$ given by

$$P_t f(x) := \int_S f(\xi) p_t(x, d\xi), \quad x \in S_\Delta,$$

and acting on all Borel-measurable functions $f : S_\Delta \rightarrow \mathbb{R}$ for which the integral is well defined. Here, $(p_t)_{t \geq 0}$ denotes the transition function, which satisfies beside the standard conditions (see [22, Definition III.1.1]) the following properties:

- (i) for all $x \in S_\Delta$, $p_0(x, \cdot) = \delta_x$, where δ_x denotes the Dirac measure;
- (ii) for all $t \geq 0$ and $x \in S$, $p_t(x, \{\Delta\}) = 1 - p_t(x, S)$ and $p_t(\Delta, \{\Delta\}) = 1$.

Since the theory of polynomial processes always deals with a Markov process $(X_t)_{t \geq 0}$ having the property that $(f(X_t))_{t \geq 0}$ is a special semimartingale for all linear functionals f on \mathbb{R}^n (extended by 0 on Δ), we can bona fide assume that the probability space Ω is the space of càdlàg functions $\omega : \mathbb{R}_+ \rightarrow S_\Delta$ such that $\omega(t-) = \Delta$ and $\omega(t) = \Delta$ implies $\omega(s) = \Delta$ for all $s \geq t$. This follows from the simple conclusion (i) to (iii) of Theorem 2.7, martingale regularity, and from the remarks at the bottom of Ref. [22, p. 245]. We thus understand X as the coordinate process $X_t(\omega) = \omega(t)$ and denote by (\mathcal{F}_t^0) the filtration generated by X and set $\mathcal{F}^0 = \bigvee_{t \geq 0} \mathcal{F}_t^0$.

In the sequel, we consider some right-continuous filtration $(\mathcal{F}_t)_{t \geq 0}$ satisfying $\mathcal{F}_t^0 \subseteq \mathcal{F}_t$ and $\mathcal{F} = \bigvee_{t \geq 0} \mathcal{F}_t$. We finally assume that for each $x \in S_\Delta$, there exists a probability measure \mathbb{P}_x on (Ω, \mathcal{F}) such that X is Markovian relative to (\mathcal{F}_t) with semigroup (P_t) , that is,

$$\mathbb{E}_x[f(X_{t+s}) \mid \mathcal{F}_s] = \mathbb{E}_{X_s}[f(X_t)] = P_t f(X_s) \quad \mathbb{P}_x\text{-a.s.}$$

for all $x \in S_\Delta$, $s, t \in [0, \infty)$ and all Borel-measurable functions $f : S_\Delta \rightarrow \mathbb{R}$ satisfying $\mathbb{E}_x[|f(X_t)|] < \infty$ for all $t \geq 0$ and $x \in S$.

2.1 Definition and characterization of polynomial processes

For the treatment of polynomial processes, we have chosen a framework of stochastic analysis leading to easy-to-verify conditions for a process to be m -polynomial (see Theorems 2.10 and 2.15). Before giving the precise definition of polynomial processes, let us introduce some notation.

Let \mathcal{P}_m denote the finite-dimensional vector space of polynomials up to degree $m \geq 0$ on S , i.e., the restriction of polynomials on \mathbb{R}^n to S , defined by

$$\mathcal{P}_m := \left\{ S \ni x \mapsto \sum_{|\mathbf{k}|=0}^m \alpha_{\mathbf{k}} x^{\mathbf{k}}, \Delta \mapsto 0 \mid \alpha_{\mathbf{k}} \in \mathbb{R} \right\},$$

where we use multi-index notation $\mathbf{k} = (k_1, \dots, k_n) \in \mathbb{N}_0^n$, $|\mathbf{k}| = k_1 + \dots + k_n$ and $x^{\mathbf{k}} = x_1^{k_1} \dots x_n^{k_n}$. The dimension of \mathcal{P}_m is denoted by $N < \infty$ and depends on S : if S is a single point, the dimension is always 1 and if S is the whole space \mathbb{R}^n , it is maximal.

Moreover, for every multi-index \mathbf{k} , we define functions $f_{\mathbf{k}}$ by setting

$$f_{\mathbf{k}}(x) = \begin{cases} x^{\mathbf{k}} & \text{if } x \in S, \\ 0 & \text{if } x = \Delta. \end{cases} \quad (2.1)$$

Furthermore, we write f_i when $\mathbf{k} = e_i$ and f_{ij} when $\mathbf{k} = e_i + e_j$, where e_i , for $i \in \{1, \dots, n\}$, denotes the i th canonical basis vector. Then the space \mathcal{P}_m clearly corresponds to the linear hull of the functions $\{f_{\mathbf{k}}, |\mathbf{k}| \leq m\}$.

Here is our main definition.

Definition 2.1 We call an S_Δ -valued time-homogeneous Markov process *m-polynomial* if we have for all $k \in \{0, \dots, m\}$, all $f \in \mathcal{P}_k$, $x \in S$ and $t \geq 0$ that

$$x \mapsto P_t f(x) \in \mathcal{P}_k.$$

Additionally, we assume that $t \mapsto P_t f(x)$ is continuous at $t = 0$ for all $f \in \mathcal{P}_m$. If X is *m-polynomial* for all $m \geq 0$, then it is called *polynomial*.

Remark 2.2 (i) Let us stress that in the above definition it is implicitly assumed that

$$P_t |f|(x) = \mathbb{E}_x[|f(X_t)|] = \int_S |f(\xi)| p_t(x, d\xi) < \infty$$

for every $f \in \mathcal{P}_m$, $x \in S_\Delta$ and $t \geq 0$; otherwise the expression $P_t f(x) = \mathbb{E}_x[f(X_t)]$ would not even be well defined.

(ii) The subtlety of Definition 2.1 lies in the fact that we assume $P_t(\mathcal{P}_k) \subset \mathcal{P}_k$ for all $k \in \{0, \dots, m\}$ (compare with Remark 2.11(iv)). The assumption that $P_t(\mathcal{P}_m) \subset \mathcal{P}_m$ only for m , but not for smaller degrees, is not sufficient for our proofs of Theorems 2.10 and 2.15, which we consider as the most important assertions from the point of view of applications.

Let us introduce the following notion of an extended Markov generator, which is due to Dynkin (see e.g. [6, Definition 7.1]) and which we shall use to characterize *m-polynomial* processes.

Definition 2.3 An operator \mathcal{G} with domain $\mathcal{D}_{\mathcal{G}}$ is called *extended generator* for some Markov process X if $\mathcal{D}_{\mathcal{G}}$ consists of those Borel-measurable functions $f : S_\Delta \rightarrow \mathbb{R}$ for which there exists a function $\mathcal{G}f$ such that the process

$$M_t^f := f(X_t) - f(x) - \int_0^t \mathcal{G}f(X_s) ds \quad (2.2)$$

is well defined and an $(\mathcal{F}_t, \mathbb{P}_x)$ -local martingale for every $x \in S_\Delta$.

Remark 2.4 We define the lifetime of the process by

$$T_\Delta(\omega) = \inf\{t \mid X_t(\omega) = \Delta\},$$

where we set $\inf \emptyset = \infty$. Since $\{T_\Delta < t\} = \bigcup_{q < t, q \in \mathbb{Q}} \{X_q = \Delta\} \in \mathcal{F}_t$ and as (\mathcal{F}_t) is supposed to be right-continuous, T_Δ is an \mathcal{F}_t -stopping time. Due to our convention $f(\Delta) = 0$, the local martingale property of (2.2) is therefore equivalent to

$$f(X_t)1_{\{t < T_\Delta\}} - f(x) - \int_0^{t \wedge T_\Delta} \mathcal{G}f(X_s) ds$$

being a local martingale.

Remark 2.5 Suppose that f lies in the domain of the extended generator and satisfies $P_t|f|(x) < \infty$ for all $t \geq 0$ and $x \in S$. Then M^f as defined in (2.2) is a true martingale if and only if all increments of $f(X_t) - f(x) - \int_0^t \mathcal{G}f(X_s) ds$ have vanishing expectation, i.e., for all $u < t$,

$$\mathbb{E}_x \left[f(X_t) - f(X_u) - \int_u^t \mathcal{G}f(X_s) ds \right] = P_t f(x) - P_u f(x) - \int_u^t P_s \mathcal{G}f(x) ds = 0.$$

In particular, by Fubini's theorem, $\int_0^t P_s \mathcal{G}f(x) ds$ exists on finite time intervals and thus also $P_s |\mathcal{G}f|(x)$ for almost all s with respect to Lebesgue measure.

The following lemma is well known for Feller processes (but perhaps not in our particular setting) and aims to establish a connection between the Kolmogorov backward equation and the extended generator introduced in (2.3).

Lemma 2.6 Let X be a time-homogeneous Markov process with semigroup (P_t) and denote by $f : S_\Delta \rightarrow \mathbb{R}$ some function satisfying $P_t|f|(x) < \infty$ for all $t \geq 0$ and $x \in S$. If f lies in the domain of the extended generator, $f \in \mathcal{D}_{\mathcal{G}}$, and if M^f as defined in (2.2) is additionally a true martingale, then we have:

- (i) For any given $s \geq 0$, $M^{P_s f}$ is a true martingale, in particular $P_s f \in \mathcal{D}_{\mathcal{G}}$, and $\mathcal{G}P_s f = P_s \mathcal{G}f$.
- (ii) If $t \mapsto P_t \mathcal{G}f(x)$ is continuous at $t = 0$, then $P_t f$ solves the Kolmogorov backward equation, that is,

$$\frac{\partial u(t, x)}{\partial t} = \mathcal{G}u(t, x), \quad u(0, x) = f(x).$$

Proof For the first statement, we show that

$$P_s f(X_t) - P_s f(x) - \int_0^t P_s \mathcal{G}f(X_r) dr$$

is a true $(\mathcal{F}_t, \mathbb{P}_x)$ -martingale for any fixed $s \geq 0$. By the definition of the extended generator, this then implies that $P_s f \in \mathcal{D}_{\mathcal{G}}$ and $\mathcal{G}P_s f = P_s \mathcal{G}f$. Indeed, we have by the assumption $P_t|f|(x) < \infty$ for all $t \geq 0$ and $x \in S$ and Remark 2.5 that $f(X_t)$ and $\mathcal{G}f(X_t)$ are integrable for every $t \geq 0$, hence $P_s f(X_t)$ and $P_s \mathcal{G}f(X_t)$ as well. Therefore the following expectation is well defined and we obtain for $u \leq t$ that

$$\begin{aligned} & \mathbb{E}_x \left[P_s f(X_t) - P_s f(x) - \int_0^t P_s \mathcal{G}f(X_r) dr \mid \mathcal{F}_u \right] \\ &= P_s f(X_u) - P_s f(x) - \int_0^u P_s \mathcal{G}f(X_r) dr \\ &+ \mathbb{E}_x \left[P_s f(X_t) - P_s f(X_u) - \int_u^t P_s \mathcal{G}f(X_r) dr \mid \mathcal{F}_u \right]. \end{aligned}$$

By the Markov property, the conditional expectation on the right is equal to

$$\mathbb{E}_{X_u} \left[P_s f(X_{t-u}) - P_s f(X_0) - \int_0^{t-u} P_s \mathcal{G} f(X_r) dr \right].$$

But for any $y \in S_\Delta$, we have

$$\begin{aligned} & \mathbb{E}_y \left[P_s f(X_{t-u}) - P_s f(X_0) - \int_0^{t-u} P_s \mathcal{G} f(X_r) dr \right] \\ &= P_{s+t-u} f(y) - P_s f(y) - \int_s^{s+t-u} P_r \mathcal{G} f(y) d\tilde{r} \\ &= 0, \end{aligned}$$

where the last equality follows from Remark 2.5. This completes the proof of (i).

Statement (ii) follows from Remark 2.5, the continuity of $t \mapsto P_t \mathcal{G} f(x)$ and from assertion (i), since

$$\begin{aligned} \frac{\partial P_t f(x)}{\partial t} &= \lim_{h \rightarrow 0} \frac{P_{t+h} f(x) - P_t f(x)}{h} = \lim_{h \rightarrow 0} P_t \frac{P_h f(x) - f(x)}{h} \\ &= \lim_{h \rightarrow 0} P_t \frac{1}{h} \int_0^h P_s \mathcal{G} f(x) ds = P_t \mathcal{G} f(x) = \mathcal{G} P_t f(x). \quad \square \end{aligned}$$

Let us now state our first theorem which is a consequence of elementary results in semigroup theory.

Theorem 2.7 *Let X be a time-homogeneous Markov process with state space S_Δ and semigroup (P_t) . Then the following assertions are equivalent:*

- (i) X is m -polynomial for some $m \geq 0$.
- (ii) For every $k \in \{0, \dots, m\}$, there exists a linear map A on \mathcal{P}_k such that for all $t \geq 0$, (P_t) restricted to \mathcal{P}_k can be written as

$$P_t|_{\mathcal{P}_k} = e^{tA}.$$

- (iii) For all $f \in \mathcal{P}_m$, $x \in S_\Delta$ and $t \geq 0$, $P_t|f|(x) = \mathbb{E}_x[|f(X_t)|] < \infty$ and

$$M_t^f = f(X_t) - f(x) - \int_0^t \mathcal{G} f(X_s) ds$$

is a true $(\mathcal{F}_t, \mathbb{P}_x)$ -martingale, and the extended generator \mathcal{G} satisfies $\mathcal{G}(\mathcal{P}_k) \subset \mathcal{P}_k$ for all $k \in \{0, 1, \dots, m\}$.

Proof To prove the above equivalences, we show (i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (i).

Throughout the proof, let $k \in \{0, \dots, m\}$ be fixed. We start by showing that (i) \Rightarrow (ii). By the definition of an m -polynomial process, the Markovian semigroup (P_t) induces a semigroup of operators on \mathcal{P}_k . Since \mathcal{P}_k is finite-dimensional and $t \mapsto P_t|_{\mathcal{P}_k}$ is continuous at $t = 0$, standard results of semigroup theory (see e.g.

[9, Theorem 2.9]) imply the representation of $P_t|_{\mathcal{P}_k}$ as matrix exponential, that is, there exists some linear map A on \mathcal{P}_k such that $P_t|_{\mathcal{P}_k} = e^{tA}$.

Next, we show (ii) \Rightarrow (iii). For every $f \in \mathcal{P}_k$ we have by (ii) that $Af \in \mathcal{P}_k$ and

$$P_t f - f - \int_0^t P_s A f \, ds = e^{tA} f - f - \int_0^t e^{sA} A f \, ds = 0.$$

Thus, $f(X_t) - f(x) - \int_0^t A f(X_s) \, ds$ is an $(\mathcal{F}_t, \mathbb{P}_x)$ -martingale. Hence f lies in $D_{\mathcal{G}}$ and $\mathcal{G}f = Af$, implying that $\mathcal{G}(\mathcal{P}_k) \subset \mathcal{P}_k$ holds true.

In order to prove (iii) \Rightarrow (i), we consider the Kolmogorov backward equation for an initial value $u(0, \cdot) = f \in \mathcal{P}_k$, i.e.,

$$\frac{\partial u(t, x)}{\partial t} = \mathcal{G}u(t, x).$$

By Lemma 2.6(ii), $P_t f$ solves the Kolmogorov equation since $t \mapsto P_t \mathcal{G}f(x)$ is continuous at $t = 0$ for any $f \in \mathcal{P}_k$. This follows from the fact that \mathcal{G} maps \mathcal{P}_k to itself and from the martingale property of M^f , which implies

$$P_t f(x) = f(x) + \int_0^t P_s \mathcal{G}f(x) \, ds,$$

thus in particular continuity of $t \mapsto P_t f(x)$ for any $f \in \mathcal{P}_k$. By choosing a basis $\{e_1, \dots, e_N\}$ of \mathcal{P}_k , we can define a linear map A on \mathcal{P}_k by setting

$$\mathcal{G}e_i =: \sum_{j=1}^N A_{ij} e_j.$$

Then $\mathcal{G}|_{\mathcal{P}_k} = A$. On \mathcal{P}_k , the Kolmogorov backward equation thus reduces to the linear ODE

$$\frac{\partial u(t)}{\partial t} = Au(t), \quad u(0) = f,$$

whose unique solution is given by $e^{tA} f$ (see e.g. [9, Theorem 2.8]). Hence on \mathcal{P}_k , $P_t f$ is equal to $e^{tA} f$ and is therefore a polynomial of degree smaller than or equal to k . Since this holds true for all $k \in \{0, 1, \dots, m\}$, X is m -polynomial. \square

Remark 2.8 There is no need in assertion (ii) to restrict the time parameter t to \mathbb{R}_+ since for $t \in \mathbb{R}$, (e^{tA}) extends to a group.

The equivalence (iii) \Leftrightarrow (i) of Theorem 2.7 provides a characterization of m -polynomial processes in terms of the extended generator, but under the additional assumption that M^f as defined in (2.2) is a *true* martingale. If $m \geq 2$ is an even number, this latter condition is no longer needed, since the local martingales M^f turn out to be always true martingales. In other words, for even numbers $m \geq 2$, Condition 2.9 below is necessary and sufficient for X to be an m -polynomial process.

Condition 2.9 \mathcal{P}_m lies in the domain of the extended generator, i.e., for all $f \in \mathcal{P}_m$, $x \in S_\Delta$ and $t \geq 0$, M^f as defined in (2.2) is an $(\mathcal{F}_t, \mathbb{P}_x)$ -local martingale, and $\mathcal{G}(\mathcal{P}_k) \subset \mathcal{P}_k$ for all $k \in \{0, 1, \dots, m\}$.

Theorem 2.10 Let X be a time-homogeneous Markov process with state space S_Δ and let $m \geq 2$ be an even number. Then X is an m -polynomial process if and only if Condition 2.9 is satisfied.

Proof The necessary direction, that is, Condition 2.9 holds if X is an m -polynomial process, is an obvious implication of Theorem 2.7 (i) \Rightarrow (iii) (of course for all m , not only even numbers larger than or equal to 2).

For the sufficient direction we prove, for every $f \in \mathcal{P}_m$, $x \in S_\Delta$ and $t \geq 0$, that $P_t|f|(x) = \mathbb{E}_x[|f(X_t)|] < \infty$ and that (M_t^f) is a true $(\mathcal{F}_t, \mathbb{P}_x)$ -martingale. Then Theorem 2.7 (iii) \Rightarrow (i) yields the assertion.

First, fix some $T > 0$ and some increasing sequence $(T_j)_{j \in \mathbb{N}}$ of stopping times with $\lim_{j \rightarrow \infty} T_j = \infty$ \mathbb{P}_x -a.s. such that $(M_{t \wedge T_j}^f)_{t \geq 0}$ are martingales for all $f \in \mathcal{P}_m$. Furthermore we set

$$F(x) := 1 + \sum_{i=1}^n f_i(x)^m,$$

where f_i , $i \in \{1, \dots, n\}$, are given by (2.1). We notice that there is some finite constant K such that

$$|\mathcal{G}f(x)| \leq KF(x)$$

for all $x \in S_\Delta$. Hence we obtain for $0 \leq t \leq T$ and $x \in S_\Delta$ that

$$\begin{aligned} \mathbb{E}_x[F(X_{t \wedge T_j})] &= F(x) + \mathbb{E}_x \left[\int_0^{t \wedge T_j} \mathcal{G}F(X_u) du \right] \\ &\leq F(x) + K \mathbb{E}_x \left[\int_0^{t \wedge T_j} F(X_u) du \right] \\ &\leq F(x) + K \mathbb{E}_x \left[\int_0^t F(X_{u \wedge T_j}) du \right]. \end{aligned}$$

Since the stopping times (T_j) can be chosen such that $X_{u \wedge T_j-}$ is bounded, the right-hand side of the above inequality is finite and Gronwall's lemma yields

$$\mathbb{E}_x[F(X_{t \wedge T_j})] \leq F(x)e^{Kt} \quad (2.3)$$

for all $0 \leq t \leq T$, $j \in \mathbb{N}$ and $x \in S_\Delta$. Due to the nonnegativity of F , we have by Fatou's lemma

$$\mathbb{E}_x[F(X_t)] = \mathbb{E}_x \left[\lim_{j \rightarrow \infty} F(X_{t \wedge T_j}) \right] \leq \liminf_{j \rightarrow \infty} \mathbb{E}_x[F(X_{t \wedge T_j})] \leq F(x)e^{Kt} \quad (2.4)$$

for all $0 \leq t \leq T$. Hence $P_t|f|(x) = \mathbb{E}_x[|f(X_t)|] < \infty$ for $f \in \mathcal{P}_m$.

Next we show for each $f \in \mathcal{P}_m$ and $x \in S_\Delta$ that

$$\mathbb{E}_x \left[\sup_{t \leq T} |M_t^f| \right] < \infty. \quad (2.5)$$

First, let $f \in \mathcal{P}_k$ for $k < m$ be fixed and set $p = m/k$. Notice that there is a finite constant \tilde{K} such that

$$|f(x)|^p \leq \tilde{K} F(x) \quad \text{and} \quad |\mathcal{G}f(x)|^p \leq \tilde{K} F(x)$$

for all $x \in S_\Delta$. We then have

$$\begin{aligned} |M_{t \wedge T_j}^f|^p &= \left| f(X_{t \wedge T_j}) - f(x) - \int_0^{t \wedge T_j} \mathcal{G}f(X_u) du \right|^p \\ &\leq C \left(F(X_{t \wedge T_j}) + F(x) + \int_0^{t \wedge T_j} F(X_u) du \right) \\ &\leq C \left(F(X_{t \wedge T_j}) + F(x) + \int_0^t F(X_{u \wedge T_j}) du \right), \end{aligned}$$

for appropriate positive constants C and $t \leq T$. Taking expectations and using (2.3), we see that for each fixed x there exists some finite constant C_x such that

$$\mathbb{E}_x[|M_{t \wedge T_j}^f|^p] \leq C_x,$$

for all $j \in \mathbb{N}$ and $t \leq T$. Moreover, by Doob's maximal L^p -inequality for $p > 1$, we have for all j that $\mathbb{E}_x[\sup_{t \leq T} |M_{t \wedge T_j}^f|^p] \leq C \mathbb{E}_x[|M_{T \wedge T_j}^f|^p] \leq C_x$. Since the left-hand side is increasing in j , monotone convergence yields (2.5) for this f , in particular

$$\sup_{t \leq T} |M_t^f| \in L^p. \quad (2.6)$$

Finally let us deal with the case $k = m$. We consider the polynomial $f(x) = f_i(x)^q$ for $i \in \{1, \dots, n\}$ and $q = m/2$, which is an integer by hypothesis. For notational convenience, we write $N = M^f$ and estimate again

$$\begin{aligned} f(X_t)^2 &= \left(N_t + f(x) + \int_0^t \mathcal{G}f(X_u) du \right)^2 \\ &\leq C \left(N_t^2 + f(x)^2 + \int_0^t |\mathcal{G}f(X_u)|^2 du \right) \\ &\leq C \left(N_t^2 + f(x)^2 + \int_0^t F(X_u) du \right). \end{aligned}$$

Therefore,

$$\sup_{t \leq T} f(X_t)^2 \leq C \left(\sup_{t \leq T} N_t^2 + f(x)^2 + \int_0^T F(X_u) du \right). \quad (2.7)$$

Due to (2.6), $\sup_{t \leq T} |N_t| \in L^2$ and by (2.4), we have $\mathbb{E}_x[\int_0^T F(X_u) du] < \infty$. Hence we conclude that the right-hand side of (2.7) is integrable. Summing over all i yields integrability of $\sup_{t \leq T} F(X_t)$, which implies (2.5) for all $f \in \mathcal{P}_m$. \square

Remark 2.11 (i) If X is an m -polynomial process, then $Z = (X, X^2, \dots, X^m)$ is a 1-polynomial process. If m is even, the analysis of m -polynomial processes could be reduced to the study of 2-polynomial processes at the cost of a more complicated state space, due to the construction $Z' = (X, X^2, \dots, X^{\frac{m}{2}})$.

(ii) Let us remark that the condition $m \geq 2$ in Theorem 2.10 is necessary for (2.2) to be a true martingale. Indeed, the inverse 3-dimensional Bessel process defined by $X = \frac{1}{\|B\|}$, where B denotes a 3-dimensional Brownian motion started at $B_0 \neq 0$, satisfies

$$dX_t = -X_t^2 dW_t, \quad X_0 = x = \frac{1}{\|B_0\|},$$

where W is a one-dimensional standard Brownian motion. The extended generator is therefore given by

$$\mathcal{G}f(x) = \frac{1}{2}x^4 \frac{d^2 f(x)}{dx^2}.$$

Hence $\mathcal{G}(\mathcal{P}_1) = 0$, but

$$X_t - x - \int_0^t \mathcal{G}X_s ds = X_t - x$$

is a strict local martingale, so X is not a 1-polynomial process.

(iii) In Definition 2.1, we require m -polynomial processes to be also k -polynomial for all $k \in \{0, \dots, m\}$, that is, we implicitly exclude processes whose extended generator maps polynomials of degree $k < m$ to polynomials of degree greater than $k \leq m$, while $\mathcal{G}(\mathcal{P}_m) \subset \mathcal{P}_m$ still holds true. Consider for instance

$$dX_t = \left(\frac{1}{2} - bX_t + \frac{1}{2}X_t^2 \right) dt + \sqrt{X_t^2(1 - X_t)} dW_t, \quad X_0 = x \in [0, 1],$$

where $b \geq 1$ and W is a one-dimensional standard Brownian motion. The state space is the interval $S = [0, 1]$ and we have

$$\mathcal{G}f(x) = \left(\frac{1}{2} - bx + \frac{1}{2}x^2 \right) \frac{df(x)}{dx} + \frac{1}{2}x^2(1 - x) \frac{d^2 f(x)}{dx^2}.$$

Thus $\mathcal{G}(\mathcal{P}_1) \subset \mathcal{P}_2$, while $\mathcal{G}(\mathcal{P}_2) \subset \mathcal{P}_2$. Due to compactness of the state space, it follows that M^f is a true martingale for $f \in \mathcal{P}_2$ and hence we can conclude that $P_t(\mathcal{P}_2) \subset \mathcal{P}_2$, but $P_t(\mathcal{P}_1) \not\subset \mathcal{P}_1$ as in the proof of Theorem 2.7 (iii) \Rightarrow (i). On the other hand—even though all moments of X exist and $\mathcal{G}(\mathcal{P}_2) \subset \mathcal{P}_2$ —the subspace $P_t(\mathcal{P}_3)$ does not only consist of polynomials any more.

2.2 Polynomial processes and semimartingales

The purpose of this section is to characterize polynomial processes as special semimartingales with characteristics of a particular form. Indeed, in Proposition 2.12 below, we prove the equivalence of Condition 2.9 and the fact that the process $(X_t 1_{\{t < T_\Delta\}})$ is a special semimartingale whose characteristics are essentially polynomials (of a particular degree) in X and absolutely continuous with respect to Lebesgue measure.

Having established the particular form of the semimartingale characteristics, the characterization of the extended generator then simply follows from Itô's formula and is given by (2.13). This in turn allows us to state easy-to-verify conditions which guarantee that X is an m -polynomial process for all $m \geq 2$. Notice that this fills a gap which is left open in Theorem 2.10, where only the case of even m is considered.

In order to formulate the following proposition concisely, we set

$$Y_t := X_t 1_{\{t < T_\Delta\}} = (f_1(X_t), \dots, f_n(X_t))^T$$

and define C_m^2 as the space of twice continuously differentiable functions $g : S \rightarrow \mathbb{R}$ for which there exists some constant \tilde{C} such that

$$|g(x)| + \sum_{i=1}^n |D_i g(x)| + \sum_{i,j=1}^n |D_{ij} g(x)| \leq \tilde{C}(1 + \|x\|^m).$$

Proposition 2.12 *Let X be a time-homogeneous Markov process with state space S_Δ and let $m \geq 2$. Then the following assertions are equivalent:*

- (i) *Condition 2.9 holds, i.e., $\mathcal{P}_m \subset \mathcal{D}_{\mathcal{G}}$ and $\mathcal{G}(\mathcal{P}_k) \subset \mathcal{P}_k$ for all $k \in \{0, 1, \dots, m\}$.*
- (ii) *$(Y_t) = (X_t 1_{\{t < T_\Delta\}})$ is a semimartingale with respect to the stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P}_x)$. Moreover,*

$$\mathbb{P}_x[t < T_\Delta] = e^{-\gamma t} \tag{2.8}$$

for some constant $\gamma \geq 0$, and the semimartingale characteristics (B, C, ν) associated with the “truncation function” $\chi(\xi) = \xi$ satisfy¹

$$B_{t,i} = \int_0^t b_i(X_s) ds, \tag{2.9}$$

$$C_{t,ij} + \int_0^t \int_{\mathbb{R}^n} \xi_i \xi_j \nu(ds, d\xi) = \int_0^t a_{ij}(X_s) ds, \tag{2.10}$$

where $b_i \in \mathcal{P}_1$ and $a_{ij} \in \mathcal{P}_2$. Furthermore, the characteristics C and ν can be written as

$$C_{t,ij} = \int_0^t c_{ij}(X_s) ds, \quad \nu(\omega; dt, d\xi) = K(X_t(\omega), d\xi) dt, \tag{2.11}$$

¹ All statements concerning the characteristics are meant up to an evanescent set.

where c denotes some Borel-measurable function taking values in the set of positive semidefinite matrices and K is a positive kernel from (S, \mathcal{S}) into $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ satisfying $K(x, \{0\}) = 0$ and $\int_{\{\|\xi\| \leq 1\}} \|\xi\|^2 K(x, d\xi) < \infty$. Finally, we have for all $|\mathbf{k}| \in \{3, \dots, m\}$ that

$$\int_{\mathbb{R}^n} \xi^{\mathbf{k}} K(x, d\xi) = \sum_{|\mathbf{j}|=0}^{|\mathbf{k}|} \alpha_{\mathbf{j}} f^{\mathbf{j}}(x), \quad (2.12)$$

where $\alpha_{\mathbf{j}}$ denote some finite coefficients.

- (iii) C_m^2 lies in the domain of the extended generator of X , and for all $g \in C_m^2$, \mathcal{G} is given by

$$\begin{aligned} \mathcal{G}g(x) &= \sum_{i=1}^n D_i g(x) (b_i(x) + \gamma f_i(x)) + \frac{1}{2} \sum_{i,j=1}^n D_{ij} g(x) c_{ij}(x) - \gamma g(x) \\ &\quad + \int_{\mathbb{R}^n} \left(g(x + \xi) - g(x) - \sum_{i=1}^n D_i g(x) \xi_i \right) \\ &\quad \times (K(x, d\xi) - \gamma f_0(x) \delta_{-x}(d\xi)), \end{aligned} \quad (2.13)$$

where γ and (b, c, K) satisfy the conditions of (ii) and $f_0(x) = 1 - 1_{\Delta}(x)$.

All conditions (i), (ii) and (iii) imply that Y is a special semimartingale.

Remark 2.13 (i) Concerning the implication (i) \Rightarrow (ii), note that the existence of $\int \xi^{\mathbf{k}} K(x, d\xi)$ for all $|\mathbf{k}| \in \{2, \dots, m\}$ follows from the fact that $\mathcal{G}f(x)$ is a well-defined polynomial for every $f \in \mathcal{P}_m$. In particular, it gives $\int \|\xi\|^k K(x, d\xi) < \infty$ for all $k \in \{2, \dots, m\}$.

(ii) Note that as a consequence of (2.10) and (2.12), we have for all $t \geq 0$ and $k \in \{2, \dots, 2\lfloor \frac{m}{2} \rfloor\}$ that

$$\int_{\mathbb{R}^n} \|\xi\|^k K(X_t, d\xi) \leq \tilde{C} (1 + \|Y_t\|^{2\lfloor \frac{k+1}{2} \rfloor}),$$

with some finite constant \tilde{C} .

(iii) Notice from the expression (2.13) for \mathcal{G} that the killing rate γ is implicitly included in the compensator K of the jump measure of Y and also in b due to the choice of the truncation function (compare [5, Sect. 3]).

Proof of Proposition 2.12 Let us first prove (i) \Rightarrow (ii). Due to Condition 2.9,

$$M_t^f = f(X_t) - f(x) - \int_0^t \mathcal{G}f(X_s) ds, \quad (2.14)$$

is an $(\mathcal{F}_t, \mathbb{P}_x)$ -local martingale for all $f \in \mathcal{P}_m$. As the process $\int_0^t \mathcal{G}f(X_s) ds$ is predictable, $f(X)$ is a special \mathbb{R} -valued semimartingale for all $f \in \mathcal{P}_m$. Note here that $f(X_t) \equiv f(X_t)1_{\{t < T_{\Delta}\}}$, due to the convention $f(\Delta) = 0$. In particular, $f(X_t)1_{\{t < T_{\Delta}\}}$

has càdlàg paths implying that $\lim_{t \rightarrow T_\Delta-} |f(X_t)| < \infty$ and $f(X)$ cannot explode. Choosing $f(x) = f_i(x)$ for $i = 1, \dots, n$ then implies that $(Y_t) = (X_t 1_{\{t < T_\Delta\}})$ is an (n -dimensional) special semimartingale.

Consider now (2.14) for $f = f_0$. Since $\mathcal{G}f_0 \in \mathcal{P}_0$, M^{f_0} is a true martingale and there exists some constant γ such that

$$\gamma 1_{\{t < T_\Delta\}} := -\mathcal{G}f_0(X_t).$$

Taking expectations thus yields

$$\begin{aligned} \mathbb{E}_x[1_{\{t < T_\Delta\}}] &= \mathbb{E}_x[f_0(X_t)] = 1 + \int_0^t \mathbb{E}_x[\mathcal{G}f_0(X_s)] ds \\ &= 1 - \gamma \int_0^t \mathbb{E}_x[1_{\{s < T_\Delta\}}] ds, \end{aligned}$$

which in turn implies (2.8).

Let now (B, C, ν) denote the characteristics of Y with respect to the “truncation function” $\chi(\xi) = \xi$. In order to determine their properties, we apply Itô’s formula to $f_{\mathbf{k}}(X_t)$ for $k = |\mathbf{k}| \in \{1, \dots, m\}$ and $X_{t,i} = x_i + M_t^{f_i} + B_{t,i}$ to obtain

$$\begin{aligned} f_{\mathbf{k}}(X_t) &= f_{\mathbf{k}}(x) + \int_0^t \sum_{i=1}^n D_i f_{\mathbf{k}}(X_{s-}) dM_s^{f_i} + \int_0^t \sum_{i=1}^n D_i f_{\mathbf{k}}(X_{s-}) dB_{s,i} \\ &\quad + \frac{1}{2} \int_0^t \sum_{i,j=1}^n D_{ij} f_{\mathbf{k}}(X_{s-}) dC_{s,ij} + \int_0^t \int_{\mathbb{R}^n} W(s, \xi) \mu^Y(ds, d\xi), \end{aligned} \quad (2.15)$$

where μ^Y denotes the random measure associated with the jumps of Y and

$$W(s, \xi) := \sum_{|\mathbf{j}|=2}^k \binom{\mathbf{k}}{\mathbf{j}} f_{\mathbf{k}-\mathbf{j}}(X_{s-}) \xi^{\mathbf{j}}.$$

Since M^{f_i} is a local martingale and $X_{s-,i}$ is càglàd, $(\int_0^t D_i f_{\mathbf{k}}(X_{s-}) dM_s^{f_i})$ is a local martingale, too, for all $i \in \{1, \dots, n\}$. Furthermore, the third and fourth terms on the right-hand side of (2.15) are predictable processes of finite variation, thus in particular of locally integrable variation by [15, Lemma I.3.10]. As $f_{\mathbf{k}}(X)$ is a special semimartingale, it follows from [15, Proposition I.4.22] that $\int_0^t \int_{\mathbb{R}^n} W(s, \xi) \mu^Y(ds, d\xi)$ is also of locally integrable variation, since it is of finite variation and for a special semimartingale, the finite variation part is locally integrable. Therefore, we have by [15, Proposition II.1.28] that

$$\int_0^t \int_{\mathbb{R}^n} W(s, \xi) \mu^Y(ds, d\xi) - \int_0^t \int_{\mathbb{R}^n} W(s, \xi) \nu(ds, d\xi)$$

is a local martingale. Combining thus (2.15) with (2.14) and using the unique decomposition of a special semimartingale into a local martingale and a predictable finite

variation process, we find

$$\begin{aligned} M^{f_{\mathbf{k}}} &= \int_0^t \sum_{i=1}^n D_i f_{\mathbf{k}}(X_{s-}) dM_s^{f_i} + \int_0^t \int_{\mathbb{R}^n} W(s, \xi) (\mu^Y(ds, d\xi) - \nu(ds, d\xi)) \\ &= f_{\mathbf{k}}(X_t) - f_{\mathbf{k}}(x) - \int_0^t \sum_{i=1}^n D_i f_{\mathbf{k}}(X_{s-}) dB_{s,i} \\ &\quad - \int_0^t \frac{1}{2} \sum_{i,j=1}^n D_{ij} f_{\mathbf{k}}(X_{s-}) dC_{s,ij} - \int_0^t \int_{\mathbb{R}^n} W(s, \xi) \nu(ds, d\xi). \end{aligned}$$

Therefore

$$\begin{aligned} \int_0^t \mathcal{G} f_{\mathbf{k}}(X_s) ds &= \int_0^t \sum_{i=1}^n D_i f_{\mathbf{k}}(X_s) dB_{s,i} + \int_0^t \frac{1}{2} \sum_{i,j=1}^n D_{ij} f_{\mathbf{k}}(X_{s-}) dC_{s,ij} \\ &\quad + \int_0^t \int_{\mathbb{R}^n} W(s, \xi) \nu(ds, d\xi). \end{aligned} \quad (2.16)$$

Consider now (2.16) for $|\mathbf{k}| = 1$, i.e., the polynomials $f_i(x)$, where $i \in \{1, \dots, n\}$. In this case (2.16) reads as

$$\int_0^t \mathcal{G} f_i(X_s) ds = B_{t,i}.$$

Setting $b_i(x) := \mathcal{G} f_i(x)$ therefore implies that $b_i \in \mathcal{P}_1$. Moreover, applying (2.16) to the quadratic polynomials $f(x) = f_{ij}(x)$ for $i, j \in \{1, \dots, n\}$ yields

$$C_{t,ij} + \int_0^t \int_{\mathbb{R}^n} \xi_i \xi_j \nu(ds, d\xi) = \int_0^t a_{ij}(X_s) ds \quad (2.17)$$

for some $a_{ij} \in \mathcal{P}_2$, since $\mathcal{G} f_{ij}(x)$ and $D_i f_{ij}(x) b_i(x) = f_j(x) b_i(x)$ lie in \mathcal{P}_2 . Hence we have proved (2.9) and (2.10).

In order to show (2.11), we define $A'_t(\omega) = \int_0^t \int_{\mathbb{R}^n} \|\xi\|^2 \nu(\omega; ds, d\xi)$. By the same arguments as in the proof of [15, Proposition II.2.9.b], there exists a random measure $K'(\omega, t; d\xi)$ on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ such that $\nu(\omega; dt, d\xi) = K'(\omega, t; d\xi) dA_t(\omega)$. Moreover, since

$$\sum_{i=1}^n C_{t,ii}(\omega) + A'_t(\omega) = \sum_{i=1}^n \int_0^t a_{ii}(X_s(\omega)) ds =: \int_0^t a_s(\omega) ds$$

and as C_{ii} , $i \in \{1, \dots, n\}$, and A' are nonnegative increasing processes, they are absolutely continuous with respect to Lebesgue measure. Hence, Ref. [15, Proposition I.3.13] implies the existence of predictable processes \tilde{c}_{ii} and H such that $C_{t,ii} = \int_0^t \tilde{c}_{s,ii} ds$ and $A'_t = \int_0^t H_s ds$. Then $\tilde{K}_{\omega,t}(d\xi) = H_t(\omega) K'(\omega, t; d\xi)$ is again a predictable random measure satisfying $\nu(\omega; dt, d\xi) = \tilde{K}_{\omega,t}(d\xi) dt$ almost surely.

Having constructed this kernel, (2.17) now becomes

$$C_{t,ij} = \int_0^t \left(a_{ij}(X_s) - \int_{\mathbb{R}^n} \xi_i \xi_j \tilde{K}_{\omega,s}(d\xi) \right) ds,$$

implying that C_{ij} for $i \neq j$ is also absolutely continuous with respect to Lebesgue measure and can therefore be written as $C_{t,ij} = \int_0^t \tilde{c}_{s,ij} ds$. Finally, by [6, Theorem 6.27] we can choose homogeneous versions for the processes \tilde{c} and \tilde{K} such that $\tilde{c}_t(\omega) = c(X_t(\omega))$ and $\tilde{K}_{\omega,t}(d\xi) = K(X_t(\omega), d\xi)$.

It remains to establish (2.12). To this end, notice that (2.16) can be written as

$$\begin{aligned} \mathcal{G}f_{\mathbf{k}}(x) &= \sum_{i=1}^n D_i f_{\mathbf{k}}(x) b_i(x) + \frac{1}{2} \sum_{i,j=1}^n D_{ij} f_{\mathbf{k}}(x) \left(c_{ij}(x) + \int_{\mathbb{R}^n} \xi_i \xi_j K(x, d\xi) \right) \\ &\quad + \int_{\mathbb{R}^n} \left(\sum_{|\mathbf{j}|=3}^k \binom{\mathbf{k}}{\mathbf{j}} f_{\mathbf{k}-\mathbf{j}}(x) \xi^{\mathbf{j}} \right) K(x, d\xi). \end{aligned}$$

Since $\mathcal{G}f_{\mathbf{k}}(x)$, $D_i f_{\mathbf{k}}(x) b_i(x)$ and $D_{ij} f_{\mathbf{k}}(x) (c_{ij}(x) + \int_{\mathbb{R}^n} \xi_i \xi_j K(x, d\xi))$ lie in \mathcal{P}_k for all $k = |\mathbf{k}| \leq m$, (2.12) simply follows by induction.

Let us now prove the implication (ii) \Rightarrow (iii). For notational simplicity we set

$$V(Y_s, \xi) := g(Y_s + \xi) - g(Y_s) - \sum_{i=1}^n D_i g(Y_s) \xi_i$$

for $g \in C_m^2$. Then, since $\int \|\xi\|^k K(x, d\xi) < \infty$ for all $k \in \{2, \dots, m\}$, which is a consequence of (2.10) and (2.12), we have

$$\int_{\mathbb{R}^n} |V(Y_s, \xi)| K(X_s, d\xi) \leq h(Y_s) + H(Y_s) \int_{\mathbb{R}^n} (\|\xi\|^2 \wedge \|\xi\|^m) K(X_s, d\xi) < \infty,$$

where h and H denote some positive (finite-valued) functions. Hence, the process $\int_0^\cdot \int V(Y_s, d\xi) K(X_s, d\xi) ds$ is of locally integrable variation and Itô's formula thus implies that

$$\begin{aligned} \mathcal{M}_t^g &:= g(Y_t) - g(x) - \int_0^t \sum_{i=1}^n D_i g(Y_s) b_i(X_s) ds - \int_0^t \frac{1}{2} \sum_{i,j=1}^n D_{ij} g(Y_s) c_{ij}(X_s) \\ &\quad - \int_0^t \int_{\mathbb{R}^n} V(Y_s, \xi) K(X_s, d\xi) ds \end{aligned} \quad (2.18)$$

is a local martingale (compare also [15, Theorem II.2.42]). Moreover,

$$-1_{\{T_\Delta \leq t\}} + \gamma(t \wedge T_\Delta) = 1_{\{t < T_\Delta\}} - 1 + \int_0^t \gamma 1_{\{s < T_\Delta\}} ds$$

is a martingale, since

$$\mathbb{E}_x \left[1_{\{t < T_\Delta\}} - 1 + \int_0^t \gamma 1_{\{s < T_\Delta\}} ds \right] = e^{-\gamma t} - 1 + \int_0^t \gamma e^{-\gamma s} ds = 0.$$

Denote now by $\mathcal{G}^\#$ the right-hand side of (2.13). We need to prove that

$$M_t^{\#g} := g(X_t) - g(x) - \int_0^t \mathcal{G}^\# g(X_s) ds$$

is a local martingale. By the definition of $\mathcal{G}^\#$ we have

$$M_t^{\#g} = \mathcal{M}_t^g - g(0)(1_{\{T_\Delta \leq t\}} - \gamma(t \wedge T_\Delta)),$$

where \mathcal{M}^g is given by (2.18). Since both terms on the right-hand side are local martingales, the same holds true for $M^{\#g}$.

Finally, (i) follows from (iii) since $\mathcal{P}_m \subset C_m^2$ and since \mathcal{G} applied to $f_{\mathbf{k}}$ maps \mathcal{P}_k into \mathcal{P}_k for $k \in \{0, \dots, m\}$, which is due to the assumptions on the characteristics. \square

Since every m -polynomial process satisfies Condition 2.9, the following corollary is an obvious consequence of Proposition 2.12.

Corollary 2.14 *Let X be an m -polynomial process with $m \geq 2$. Then the process $(Y_t) = (X_t 1_{\{t < T_\Delta\}})$ is a special semimartingale satisfying the conditions (2.8)–(2.12), and its extended generator is of the form (2.13).*

When m is an even number, the converse direction also follows easily from Theorem 2.10 and Proposition 2.12. The general case is treated in the subsequent theorem, where we provide sufficient conditions in terms of the compensator of the jump measure such that the converse statement holds true. Its proof relies on a maximal inequality which can be established for semimartingales whose characteristics satisfy the conditions (2.9)–(2.12). This is the subject of Lemma 2.17 below.

Theorem 2.15 *Let X be a time-homogeneous Markov process with state space S_Δ and let $m \geq 2$. Suppose that $(Y_t) = (X_t 1_{\{t < T_\Delta\}})$ is a semimartingale which satisfies the conditions (2.8)–(2.12) (with respect to the “truncation function” $\chi(\xi) = \xi$) or equivalently that $C_m^2 \subset \mathcal{D}_{\mathcal{G}}$ and that its extended generator \mathcal{G} is given by (2.13). If*

$$\mathbb{E}_x \left[\int_{\mathbb{R}^n} \|\xi\|^m K(X_t, d\xi) \right] < \infty \quad \text{for almost all } t \geq 0, \quad (2.19)$$

or if

$$\int_{\mathbb{R}^n} \|\xi\|^m K(X_t, d\xi) \leq \tilde{C}(1 + \|Y_t\|^m), \quad t \geq 0, \quad (2.20)$$

for some constant \tilde{C} , then X is an m -polynomial process.

Proof Similarly as in the proof of Theorem 2.10, it suffices to show that for each f , $x \in S_\Delta$ and every fixed $t \geq 0$, $P_t|f|(x) = \mathbb{E}_x[|f(X_t)|] < \infty$ and

$$\mathbb{E}_x \left[\sup_{s \leq t} |M_s^f| \right] < \infty.$$

Due to the assumptions (2.19) or (2.20), this follows from the moment estimate (2.22) proved in Lemma 2.17 below. \square

Remark 2.16 Let us remark that under (2.12), condition (2.20) is always satisfied when m is an even number (see also Remark 2.13(ii)). In this case $\mathbb{E}_x[\sup_{s \leq t} |M_s^f|]$ is always finite, as already shown in the proof of Theorem 2.10. If $m > 2$ is an odd number, Condition 2.9 together with (2.19) or (2.20) is sufficient for X to be an m -polynomial process.

Using the structure of the semimartingale characteristics derived in Proposition 2.12, we now state the announced maximal inequality and some associated moment estimates. This result is probably known, but the proof is included for convenience. Similar L^p estimates for the case of Lévy-driven SDEs can be found in [16] and [18]. In the case when jumps are described by

$$\int_0^t \int_{\mathbb{R}^d} \delta(s, \xi) N(ds, d\xi),$$

where δ denotes some predictable process and N a compensated Poisson random measure, our result could be deduced from [18, Theorem 2.11].²

Lemma 2.17 Fix $t > 0$ and let $m \geq 2$. Let Y be a semimartingale with respect to $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P}_x)$ whose characteristics (B, C, ν) associated with the “truncation function” $\chi(\xi) = \xi$ satisfy the conditions (2.9)–(2.11) given in Proposition 2.12. Then there exists a constant \tilde{C} such that

$$\begin{aligned} \mathbb{E}_x \left[\sup_{s \leq t} \|Y_s\|^m \right] &\leq \tilde{C} \left(\|x\|^m + 1 + \int_0^t \mathbb{E}_x \left[\int_{\mathbb{R}^n} \|\xi\|^m K(X_s, d\xi) \right] ds \right. \\ &\quad \left. + \int_0^t \mathbb{E}_x [\|Y_s\|^m] ds \right). \end{aligned} \quad (2.21)$$

In particular, if one of the conditions (2.19) or (2.20) is satisfied, then there exist finite constants K and \tilde{C} such that

$$\mathbb{E}_x \left[\sup_{s \leq t} \|Y_s\|^m \right] \leq K e^{\tilde{C}t}. \quad (2.22)$$

Proof For notational simplicity we only consider the case where Y is one-dimensional and thus omit all indices. Due to Proposition 2.12 and the assumptions on the characteristics, Y is a special semimartingale and its canonical decomposition is given by $Y = x + M + \int_0^\cdot b(X_u) du$, where we write M for M^{f_1} . Denote by Z the quadratic variation of the purely discontinuous martingale part of Y , that is,

$$Z_t = \sum_{s \leq t} (\Delta Y_s)^2 = \int_0^t \int_{\mathbb{R}} \xi^2 \mu^Y(ds, d\xi),$$

²We thank Martin Schweizer for pointing out this result to us.

where μ^Y is the random measure associated with the jumps of Y . Define furthermore stopping times

$$\begin{aligned} T_j^Y &= \inf\{t \geq 0 \mid |Y_t| \geq j \text{ or } |Y_{t-}| \geq j\}, \\ T_j^Z &= \inf\{t \geq 0 \mid |Z_t| \geq j \text{ or } |Z_{t-}| \geq j\} \end{aligned}$$

and set $T_j = T_j^Y \wedge T_j^Z$. We can estimate

$$\sup_{s \leq t} |Y_s^{T_j}|^m \leq \tilde{C} \left(|x|^m + \sup_{s \leq t} |M_s^{T_j}|^m + \sup_{s \leq t} \left| \int_0^{s \wedge T_j} b(X_u) du \right|^m \right),$$

where \tilde{C} denotes some constant which may vary from line to line.

Since $b \in \mathcal{P}_1$ we have

$$\sup_{s \leq t} \left| \int_0^{s \wedge T_j} b(X_u) du \right|^m \leq \tilde{C} \left(1 + \int_0^t |Y_u^{T_j}|^m du \right). \quad (2.23)$$

Concerning $\sup_{s \leq t} |M_s^{T_j}|$, an application of the Burkholder–Davis–Gundy inequality yields

$$\mathbb{E}_x \left[\sup_{s \leq t} |M_s^{T_j}|^m \right] \leq \tilde{C} \mathbb{E}_x [[M, M]_{t \wedge T_j}^{\frac{m}{2}}] \leq \tilde{C} \mathbb{E}_x [C_{t \wedge T_j}^{\frac{m}{2}} + Z_{t \wedge T_j}^{\frac{m}{2}}]. \quad (2.24)$$

As C satisfies (2.10), we can estimate it by $C_t \leq \int_0^t a(X_s) ds$, where $a \in \mathcal{P}_2$ is non-negative, and we get

$$\mathbb{E}_x [C_{t \wedge T_j}^{\frac{m}{2}}] \leq \tilde{C} \left(1 + \int_0^t \mathbb{E}_x [|Y_s^{T_j}|^m] ds \right).$$

Therefore it remains to handle $Z_{t \wedge T_j}^{\frac{m}{2}}$. Following the approach of [16], we can write

$$\begin{aligned} Z_{t \wedge T_j}^{\frac{m}{2}} &= \sum_{s \leq t \wedge T_j} (Z_{s-} + \Delta Z_s)^{\frac{m}{2}} - (Z_{s-})^{\frac{m}{2}} \\ &= \int_0^{t \wedge T_j} \int_{\mathbb{R}^n} \left((Z_{s-} + \xi^2)^{\frac{m}{2}} - (Z_{s-})^{\frac{m}{2}} \right) \mu^Y(ds, d\xi), \end{aligned}$$

due to the fact that Z is purely discontinuous, nondecreasing and $\Delta Z_s = |\Delta Y_s|^2$. Furthermore, since ν is the predictable compensator of μ^Y , we have

$$\mathbb{E}_x [Z_{t \wedge T_j}^{\frac{m}{2}}] = \mathbb{E}_x \left[\int_0^{t \wedge T_j} \int_{\mathbb{R}^n} \left((Z_{s-} + \xi^2)^{\frac{m}{2}} - (Z_{s-})^{\frac{m}{2}} \right) \nu(ds, d\xi) \right]. \quad (2.25)$$

In the sequel, we shall use the inequalities (see [16])

$$(z + x)^p - z^p \leq 2^{p-1} (z^{p-1} x + x^p), \quad (2.26)$$

$$z^{p-1} x \leq \varepsilon z^p + \frac{x^p}{\varepsilon^{p-1}}, \quad (2.27)$$

for $x, z \geq 0$, $\varepsilon > 0$ and $p \geq 1$. Applying (2.26), (2.25) becomes

$$\mathbb{E}_x[Z_{t \wedge T_j}^{\frac{m}{2}}] \leq \mathbb{E}_x \left[\int_0^{t \wedge T_j} \int_{\mathbb{R}^n} 2^{\frac{m}{2}-1} (Z_{s-}^{\frac{m}{2}-1} \xi^2 + |\xi|^m) v(ds, d\xi) \right].$$

For the first part, we then have due to the assumption on v that

$$\begin{aligned} & \mathbb{E}_x \left[\int_0^{t \wedge T_j} \int_{\mathbb{R}^n} 2^{\frac{m}{2}-1} Z_{s-}^{\frac{m}{2}-1} \xi^2 v(ds, d\xi) \right] \\ &= \mathbb{E}_x \left[\int_0^{t \wedge T_j} 2^{\frac{m}{2}-1} Z_s^{\frac{m}{2}-1} \left(\int_{\mathbb{R}^n} \xi^2 K(X_s, d\xi) \right) ds \right] \\ &\leq \mathbb{E}_x \left[\int_0^{t \wedge T_j} 2^{\frac{m}{2}-1} Z_s^{\frac{m}{2}-1} a(X_s) ds \right] \\ &\leq \mathbb{E}_x \left[\int_0^{t \wedge T_j} \tilde{C} \left(\varepsilon Z_s^{\frac{m}{2}} + \frac{1 + |Y_s|^m}{\varepsilon^{\frac{m}{2}-1}} \right) ds \right], \end{aligned}$$

where we use that Z is nonnegative, $a \in \mathcal{P}_2$ and (2.27). Estimating

$$\int_0^{t \wedge T_j} Z_s^{\frac{m}{2}} ds \leq j^{\frac{m}{2}} \wedge Z_{t \wedge T_j}^{\frac{m}{2}},$$

which follows from the fact that Z_s is nondecreasing and $Z_{s-} \leq j$ for $s \leq T_j$, we finally obtain

$$\begin{aligned} \mathbb{E}_x[Z_{t \wedge T_j}^{\frac{m}{2}}] &\leq \tilde{C} \varepsilon \mathbb{E}_x[j^{\frac{m}{2}} \wedge Z_{t \wedge T_n}^{\frac{m}{2}}] + \mathbb{E}_x \left[\int_0^{t \wedge T_j} \frac{\tilde{C}}{\varepsilon^{\frac{m}{2}-1}} (1 + |Y_s|^m) ds \right] \\ &\quad + \mathbb{E}_x \left[\int_0^{t \wedge T_j} \int_{\mathbb{R}^n} 2^{\frac{m}{2}-1} |\xi|^m K(X_s, d\xi) ds \right]. \end{aligned}$$

Choosing $\varepsilon = \frac{1}{2\tilde{C}}$ leads to

$$\begin{aligned} \frac{1}{2} \mathbb{E}_x[Z_{t \wedge T_j}^{\frac{m}{2}}] &\leq \mathbb{E}_x[Z_{t \wedge T_j}^{\frac{m}{2}}] - \frac{1}{2} \mathbb{E}_x[j^{\frac{m}{2}} \wedge Z_{t \wedge T_j}^{\frac{m}{2}}] \\ &\leq \tilde{C} \mathbb{E}_x \left[\int_0^t \left(1 + |Y_s^{T_j}|^m + \int_{\mathbb{R}^n} |\xi|^m K(X_{s \wedge T_j}, d\xi) \right) ds \right]. \end{aligned}$$

Combining this with the estimates (2.23) and (2.24), we find

$$\begin{aligned} \mathbb{E}_x \left[\sup_{s \leq t} |Y_s^{T_j}|^m \right] &\leq \tilde{C} \left(|x|^m + 1 + \int_0^t \mathbb{E}_x \left[\int_{\mathbb{R}} |\xi|^m K(X_{s \wedge T_j}, d\xi) \right] ds \right. \\ &\quad \left. + \int_0^t \mathbb{E}_x[|Y_s^{T_j}|^m] ds \right). \end{aligned} \quad (2.28)$$

By monotone convergence we obtain (2.21).

Concerning (2.22), note that under the conditions (2.19) or (2.20), we can deduce from (2.28) that

$$\mathbb{E}_x \left[\sup_{s \leq t} |Y_s^{T_j}|^m \right] \leq K + \tilde{C} \int_0^t \mathbb{E}_x \left[\sup_{u \leq s} |Y_u^{T_j}|^m \right] ds$$

for some finite constants K and \tilde{C} . Since the right-hand side is finite due to the conditions (2.19) or (2.20) and the estimate

$$\begin{aligned} \sup_{u \leq s} |Y_u^{T_j}|^m &\leq |x|^m + j^m + \mathbb{E}_x [|\Delta Y_{s \wedge T_j}|^m] \\ &\leq |x|^m + j^m + \mathbb{E}_x \left[\int_0^{s \wedge T_j} \int_{\mathbb{R}} |\xi|^m K(X_s, d\xi) ds \right] < \infty, \end{aligned}$$

Gronwall's lemma yields

$$\mathbb{E}_x \left[\sup_{s \leq t} |Y_s^{T_j}|^m \right] \leq K e^{\tilde{C}t}$$

for all $j \in \mathbb{N}$ and the result follows from monotone convergence. \square

Remark 2.18 (i) It is important to note that the characteristics of $Y = (X_t 1_{\{t < T_\Delta\}})$ in the above statements are always specified with respect to the “truncation function” $\chi(\xi) = \xi$. While C and ν do not depend on this choice, the characteristic B does depend on χ . So if one chooses another truncation function χ' , the difference between B and B' is given by $\int_0^t \int_{\mathbb{R}^n \setminus \{0\}} (\chi'(\xi) - \chi(\xi)) \nu(ds, d\xi)$. Thus, the requirement that C and ν are as in Theorem 2.15 and

$$\left(b(X_t) + \int_{\mathbb{R}^n \setminus \{0\}} (\chi(\xi) - \chi'(\xi)) K(X_t, d\xi) \right) = \sum_{|\mathbf{k}|=0}^1 \alpha_{\mathbf{k}} f_{\mathbf{k}}(X_t) \quad (2.29)$$

is an equivalent condition guaranteeing that X is m -polynomial.

(ii) We now give two examples of kernels $K(x, d\xi)$ which satisfy the conditions of Proposition 2.12 as long as $c_{ij} \in \mathcal{P}_2$ and b satisfies (2.29).

(a) The first one essentially requires $K(x, d\xi)$ to be a quadratic polynomial in x , that is,

$$K(x, d\xi) = \left(\frac{\mu_{00}(d\xi)}{\|\xi\|^2 \wedge 1} + \sum_{i=1}^n x_i \frac{\mu_{i0}(d\xi)}{\|\xi\|^2 \wedge 1} + \sum_{i \leq j} x_i x_j \frac{\mu_{ij}(d\xi)}{\|\xi\|^2 \wedge 1} \right),$$

where all μ_{ij} are finite signed measures on \mathbb{R}^n such that $K(x, \cdot)$ is a well-defined Lévy measure for every $x \in S$. In view of Remark 2.13, it is necessary to require

$$\int_{\|\xi\| > 1} \|\xi\|^m (\mu_{ij}^+(d\xi) + \mu_{ij}^-(d\xi)) = \int_{\|\xi\| > 1} \|\xi\|^m |\mu_{ij}(d\xi)| < \infty,$$

where μ_{ij}^+, μ_{ij}^- denotes the Jordan decomposition of μ_{ij} .

- (b) Alternatively, K can be specified as the pushforward of a Lévy measure under an affine function. Let $d \geq 1$ and let

$$p : S \times \mathbb{R}^d \rightarrow \mathbb{R}^n, \quad (x, y) \mapsto p(x, y) = p^x(y) = H(y)x + h(y),$$

be an affine function in x . Here, $H : \mathbb{R}^d \rightarrow \mathbb{R}^{n \times n}$ and $h : \mathbb{R}^d \rightarrow \mathbb{R}^n$ are assumed to be measurable. We define K then by

$$K(x, d\xi) := (p^x)_* \mu(d\xi),$$

where for each $x \in S$, $(p^x)_* \mu$ denotes the pushforward of the measure μ under the map p^x . Moreover, μ is a Lévy measure on \mathbb{R}^d integrating

$$\int_{\mathbb{R}^d \setminus \{0\}} (\|H(y)\|^k + \|h(y)\|^k) \mu(dy) \quad \text{for all } k \in \{1, \dots, m\}.$$

3 Examples

In order to apply Theorem 2.15 to the following examples, we assume $m \geq 2$.

Example 3.1 (Affine processes) Every affine process X on $S = \mathbb{R}_+^p \times \mathbb{R}^{n-p}$ is m -polynomial if the killing rate is constant and if the Lévy measures μ_i for $i \in \{0, 1, \dots, p\}$ satisfy

$$\int_{\|\xi\| > 1} \|\xi\|^m \mu_i(d\xi) < \infty. \quad (3.1)$$

For details on affine processes see e.g. [7] and [17].³

Proof Since the differential characteristics of $Y = (X_t 1_{\{t < T_\Delta\}})$ are affine functions in X and since (3.1) ensures that condition (2.20) is satisfied, that is,

$$\begin{aligned} \int_{\mathbb{R}^n} \|\xi\|^m K(X_t, d\xi) &= \int_{\mathbb{R}^n} \|\xi\|^m \left(\mu_0(d\xi) + \sum_{i=1}^p X_{t,i} \mu_i(d\xi) \right) \\ &\leq \tilde{C}(1 + \|Y_t\|) \leq \tilde{C}(1 + \|Y_t\|^m) \end{aligned}$$

for some constant \tilde{C} and $m \geq 2$, the assertion follows from Theorem 2.15. \square

Let us remark that affine processes are defined via their characteristic function, which is of the form

$$\mathbb{E}_x[e^{i\langle u, X_t \rangle}] = e^{\phi(t, iu) + \langle \psi(t, iu), x \rangle}$$

for some function ϕ and ψ . This definition then implies affine semimartingale characteristics, from which the polynomial property can easily be seen due to Theorem 2.15.

³We write here μ_0 for the constant part of the jump measure in contrast to [7], where it is denoted by m .

Note also that the explicit knowledge of ϕ and ψ is not necessary to compute the moments of an affine process. Simply the knowledge of its characteristics, which determine the linear map A given in Theorem 2.7(ii), is enough (see Sect. 4). Moreover, for affine processes, it has not been proved so far that the existence of the moments of the Lévy measures automatically implies the existence of the moments of the process itself. We can conclude this simply from the more general statement of Lemma 2.17.

Example 3.2 (Exponential Lévy models) Exponential Lévy models are of the form $X = xe^L$, where L is a Lévy process on \mathbb{R} with triplet (b, c, μ) . Under the integrability assumption $\int_{|y|>1} e^{my} \mu(dy) < \infty$, which guarantees the existence of $\mathbb{E}_x[|X_t|^m]$, exponential Lévy models are m -polynomial, since we have $\mathbb{E}_x[x^m e^{mL_t}] = x^m e^{t\psi(m)}$, where ψ denotes the cumulant generating function of the Lévy process.

Example 3.3 (Lévy-driven SDEs) Let L denote a Lévy process on \mathbb{R}^d with triplet (b, c, μ) . Suppose furthermore that V_1, \dots, V_d are affine functions, i.e., we have $V_i : S \rightarrow \mathbb{R}^n$, $x \mapsto H_i x + h_i$, where $H_i \in \mathbb{R}^{n \times n}$ and $h_i \in \mathbb{R}^n$. A process X which solves the stochastic differential equation

$$dX_t = \sum_{i=1}^d V_i(X_{t-}) dL_{t,i}, \quad X_0 = x \in S,$$

and leaves S invariant is m -polynomial if the Lévy measure satisfies the moment condition

$$\int_{\|y\|>1} \|y\|^m \mu(dy) < \infty. \quad (3.2)$$

Proof For C_m^2 -functions g and general Lipschitz-continuous functions V_1, \dots, V_d , the extended generator of X with respect to some truncation function χ is given by

$$\begin{aligned} & \left\langle \sum_{i=1}^d V_i(x) b_i, \nabla g(x) \right\rangle + \frac{1}{2} \sum_{i,j=1}^n ((V_1(x) \dots V_d(x)) c (V_1(x) \dots V_d(x))')_{ij} D_{ij} g(x) \\ & + \int \left(g \left(x + \sum_{i=1}^d V_i(x) y_i \right) - g(x) - \left\langle \sum_{i=1}^d V_i(x) \chi_i(y), \nabla g(x) \right\rangle \right) \mu(dy). \end{aligned}$$

Concerning the compensator of the jump measure $K(x, d\xi)$, this example corresponds to the situation of Remark 2.18(ii) (b) with

$$p(x, y) = H(y)x + h(y) = \sum_{i=1}^d H_i y_i x + h_i y_i.$$

Condition (2.20), that is,

$$\int_{\mathbb{R}^n} \|\xi\|^m K(X_t, d\xi) = \int_{\mathbb{R}^n} \left\| \sum_{i=1}^d H_i y_i X_t + h_i y_i \right\|^m \mu(dy) \leq \tilde{C} (1 + \|X_t\|^m)$$

for some constant \tilde{C} , is satisfied due to (3.2). Hence Theorem 2.15 yields the assertion. \square

Example 3.4 (Quadratic term structure models [4]) Consider for r a quadratic term structure model, specified as nonnegative quadratic function of a one-dimensional Ornstein–Uhlenbeck process Y , i.e.,

$$r_t = R_0 + R_1 Y_t + R_2 Y_t^2$$

for appropriate $R_i \in \mathbb{R}$. Here, Y is given by

$$dY_t = (b + \beta Y_t) dt + \sigma dW_t,$$

where W is a standard Brownian motion. The joint process $X = (Y, r)$ then satisfies the dynamics

$$\begin{aligned} \begin{pmatrix} dY_t \\ dr_t \end{pmatrix} = & \left(\begin{pmatrix} b \\ R_1 b + R_2 \sigma^2 - 2R_0 \beta \end{pmatrix} + \begin{pmatrix} \beta \\ 2R_2 b - R_1 \beta \end{pmatrix} Y_t + \begin{pmatrix} 0 \\ 2\beta \end{pmatrix} r_t \right) dt \\ & + \begin{pmatrix} \sigma \\ (R_1 + 2R_2 Y_t) \sigma \end{pmatrix} dW_t \end{aligned}$$

and is therefore clearly a polynomial process with

$$C_t = \int_0^t \left(\sigma^2 \begin{pmatrix} 1 & R_1 \\ R_1 & R_1^2 \end{pmatrix} + \sigma^2 \begin{pmatrix} 0 & 2R_2 \\ 2R_2 & 4R_1 R_2 \end{pmatrix} Y_s + \sigma^2 \begin{pmatrix} 0 & 0 \\ 0 & 4R_2^2 \end{pmatrix} Y_s^2 \right) ds.$$

Example 3.5 (Jacobi process) Another example of a polynomial process is the Jacobi process (see [14]), which is the solution of the stochastic differential equation

$$dX_t = -\beta(X_t - \theta) dt + \sigma \sqrt{X_t(1 - X_t)} dW_t, \quad X_0 = x \in [0, 1],$$

on $S = [0, 1]$, where $\theta \in [0, 1]$ and $\beta, \sigma > 0$. This example can be extended by adding jumps, where the jump times correspond to those of a Poisson process with intensity λ and the jump size is a function of the process level. Indeed, if a jump occurs, then the process is reflected at $\frac{1}{2}$ so that it remains in the interval $[0, 1]$. The extended generator is given by

$$\mathcal{G}g(x) = -\beta(x - \theta) \frac{dg(x)}{dx} + \frac{1}{2} \sigma^2 (x(1 - x)) \frac{d^2 g(x)}{dx^2} + \lambda (g(1 - x) - g(x)).$$

In terms of Remark 2.18(ii) (b), we have $p(x, y) = -2yx + y$ and $\mu(dy) = \lambda \delta_1(dy)$.

Example 3.6 (Pearson diffusions) The above Example 3.5 (without jumps) as well as Ornstein–Uhlenbeck and Cox–Ingersoll–Ross processes, all of them with mean-reverting drift, can be subsumed under the class of so-called Pearson diffusions, which are the solutions to SDEs of the form

$$dX_t = -\beta(X_t - \theta) dt + \sqrt{(a + \alpha_{10} X_t + \alpha_{11} X_t^2)} dW_t, \quad X_0 = x,$$

where $\beta > 0$ and α_{10} , α_{11} and a are specified such that the square root is well defined. In view of Theorem 2.15, it is thus obvious that these processes are polynomial. Forman and Sørensen [10] give a complete classification of the different types of the Pearson diffusion in terms of their invariant distributions.

Example 3.7 (Dunkl process) The extended generator of the so-called Dunkl process (see [8, 11]) is given by

$$\begin{aligned}\mathcal{G}g(x) &= \frac{d^2g(x)}{dx^2} + \frac{\lambda}{2x^2} \int_{\mathbb{R}} \left(g(x + \xi) - g(x) - \xi \frac{dg(x)}{dx} \right) \delta_{-2x}(d\xi) \\ &= \frac{d^2g(x)}{dx^2} + \frac{\lambda}{x} \frac{dg(x)}{dx} + \frac{\lambda(g(-x) - g(x))}{2x^2}.\end{aligned}$$

Since $K(x, d\xi) = \frac{\lambda}{2x^2} \delta_{-2x}(d\xi)$ and

$$\int_{\mathbb{R}} |\xi|^m K(X_t, d\xi) = \frac{\lambda |2X_t|^m}{2X_t^2} = 2^{m-1} \lambda |X_t|^{m-2}$$

for all $m \geq 2$, we derive from Theorem 2.15 that the Dunkl process is a polynomial process.

4 Applications

By Theorem 2.7, we know that there exists a linear map A such that moments of m -polynomial processes can simply be calculated by computing e^{tA} . Indeed, by choosing a basis $\langle e_1, \dots, e_N \rangle$ of \mathcal{P}_m , the matrix corresponding to this linear map, which we also denote by $A = (A_{k\ell})_{k,\ell=1,\dots,N}$, can be obtained through $Ae_k = \sum_{\ell=1}^N A_{k\ell} e_{\ell}$. Writing f as $\sum_{k=1}^N \alpha_k e_k$, we then have

$$P_t f = (\alpha_1, \dots, \alpha_N) e^{tA} (e_1, \dots, e_N)',$$

which means that moments of polynomial processes can be evaluated simply by computing matrix exponentials.

By means of the one-dimensional Cox–Ingersoll–Ross process

$$dX_t = (b + \beta X_t) dt + \sigma \sqrt{X_t} dW_t, \quad b, \sigma \in \mathbb{R}_+, \beta \in \mathbb{R},$$

we exemplify how moments of order m can be calculated. Its extended generator is given by

$$\mathcal{A}g(x) = \frac{1}{2} \sigma^2 x \frac{d^2g(x)}{dx^2} + (b + \beta x) \frac{dg(x)}{dx}.$$

Applying \mathcal{A} to (x^0, x^1, \dots, x^m) yields the $(m+1) \times (m+1)$ matrix

$$A = \begin{pmatrix} 0 & \dots & & & & \\ b & \beta & 0 & \dots & & \\ 0 & 2b + \sigma^2 & 2\beta & 0 & \dots & \\ 0 & 0 & 3b + 3\sigma^2 & 3\beta & 0 & \dots \\ & & & \ddots & & \\ 0 & \dots & & & mb + \frac{m(m-1)}{2}\sigma^2 & m\beta \end{pmatrix}.$$

Hence, $\mathbb{E}_x[X_t^k] = P_t x^k = (0, \dots, 1, \dots, 0) e^{tA} (x^0, \dots, x^k, \dots, x^m)'$.

Remark 4.1 (i) Note that A is a lower triangular matrix, whose eigenvalues are the diagonal elements. Since in this case they are all distinct, the matrix is diagonalizable. Of course, there are many efficient algorithms to evaluate such matrix exponentials (see for example [13, 19]).

(ii) If $n > 1$, one has to apply well-known techniques from linear algebra of polynomials, in order to enumerate efficiently a basis of \mathcal{P}_m and to exploit sparsity properties of A ; see for instance [21].

4.1 Moment estimation—generalized method of moments (GMM)

In view of this easy and fast technique of moment calculation for polynomial processes, the generalized method of moments (GMM) can be applied for parameter estimation. If we are given a stationary polynomial process X , then typical functionals applied for parameter estimation are of the form

$$f(X, \theta) = \begin{pmatrix} X_t^{n_1} X_{t+s}^{m_1} - \mathbb{E}_x[X_t^{n_1} X_{t+s}^{m_1}] \\ \vdots \\ X_t^{n_q} X_{t+s}^{m_q} - \mathbb{E}_x[X_t^{n_q} X_{t+s}^{m_q}] \end{pmatrix}, \quad n_i, m_i \in \mathbb{N}, \quad 1 \leq i \leq q,$$

where θ is the set of parameters to be estimated. This functional is indeed simple to evaluate, since $\mathbb{E}_x[X_t^n X_{t+s}^m] = \mathbb{E}_x[X_t^n \mathbb{E}_{X_t}[X_{t+s}^m]]$ can also be computed easily. The usual technology of equating time averages to expectations applies for f and leads to efficient calibration methods. In the case of one-dimensional jump-diffusions, Zhou [23] already uses this method for GMM estimation.

4.2 Model calibration

In model calibration—in contrast to estimation of parameter values of a stationary process from time series data—parameters are chosen such that derivatives' prices are best explained. Here also the polynomial structure can be very helpful. Assume a polynomial process X , where derivatives' prices are known from the market. Derivatives' prices are expectations $\mathbb{E}_x[f(X_t)]$ for sufficiently many time points $t > 0$ and sufficiently many payoffs f , such that we can estimate the curves $t \mapsto P_t g(x)$ for today's initial value x and several polynomials g . In other words, we need as many

derivatives' prices as necessary to calculate (estimate) the prices of some payoffs, which are polynomials in the underlying.

Once we have those curves $t \mapsto P_t g(x)$, it is often an easy task to read off parameter values which explain this curve. This will be worked out in a follow-up paper.

4.3 Pricing—variance reduction

The fact that moments of polynomial processes are analytically known also gives rise to new and efficient techniques for pricing and hedging.

Let X be an m -polynomial process and $G : S \rightarrow \mathbb{R}^n$ a deterministic measurable map such that the discounted price processes are given through $S_t = G(X_t)$ under a martingale measure. Typically $G = \exp$ if X are log-prices. We denote by $F = \phi(S_T)$ a bounded measurable European claim for some maturity $T > 0$, whose (discounted) price at $t \geq 0$ is given by the risk-neutral valuation formula

$$p_t^F = \mathbb{E}_x[\phi(S_T) \mid \mathcal{F}_t] = \mathbb{E}_{X_t}[(\phi \circ G)(X_T)].$$

Obviously, claims of the form

$$F = f \circ G^{-1}(S_T) \tag{4.1}$$

for $f \in \mathcal{P}_m$ are analytically tractable, since we have

$$p_t^F = \mathbb{E}_x[(f \circ G^{-1})(S_T) \mid \mathcal{F}_t] = P_{T-t} f(G^{-1}(S_t)) = e^{(T-t)A} f(G^{-1}(S_t))$$

for $0 \leq t \leq T$, where A is the previously defined linear map on \mathcal{P}_m . Although claims are in practice not of the form (4.1), the explicit knowledge of the price of polynomial claims can be used for variance reduction techniques based on control variates. Instead of using the estimator $\pi_0^F = \frac{1}{L} \sum_{i=1}^L (\phi \circ G)(X_T^i)$ in a Monte Carlo simulation, where X_T^1, \dots, X_T^L are L samples of X_T , we can use

$$\hat{\pi}_0^F = \frac{1}{L} \sum_{i=1}^L \left((\phi \circ G)(X_T^i) - (f(X_T^i) - \mathbb{E}_x[f(X_T)]) \right),$$

where $f \in \mathcal{P}_m$ is an approximation of $\phi \circ G$ and serves as control variate. Both estimators are unbiased and the second clearly outperforms the first since we have $\text{Var}(\hat{\pi}_0^F) < \text{Var}(\pi_0^F)$, where the ratio of the variances depends on the accuracy of the polynomial approximation.

It is worth mentioning that the previous pricing algorithm has also important consequences for hedging, since the Greeks for “polynomial claims” $F = f(X_T)$ can also be calculated explicitly and efficiently. The coefficients of the polynomial $x \mapsto \mathbb{E}_x[f(X_T)]$ can be computed using matrix exponentiation, and taking derivatives of this polynomial is then a simple algebraic operation. To be more precise, the sensitivities of the price process with respect to the factors X can be calculated by

$$\nabla p_t^F = \nabla P_{T-t} f(G^{-1}(S_t)) \nabla G^{-1}(S_t).$$

Assuming a complete market situation, the claim $\phi(S_T) = \phi \circ G(X_T)$ can be replicated by a trading strategy η , i.e.,

$$\phi(S_T) = \mathbb{E}_x[\phi(S_T)] + \int_0^T \eta_t dS_t.$$

Similarly, the polynomial claim $f(X_T)$ can be replicated by the delta-hedging strategy ∇p^F and we conclude that

$$\phi(S_T) - f(X_T) = \mathbb{E}_x[\phi(S_T)] - \mathbb{E}_x[f(X_T)] + \int_0^T (\eta - \nabla p_t^F) dS_t.$$

Therefore, if we assume that $\phi(S_T) - f(X_T)$ has small variance, then also the stochastic integral representing the difference of the cumulative gains and losses of the two hedging portfolios, namely the one built by the unknown strategy η and the one built by the known strategy ∇p^F , is small.

Concerning the approximation of $\phi \circ G$ by a polynomial, let us consider the case $S_t = G(X_{t,1})$ with $G: \mathbb{R} \rightarrow \mathbb{R}_+$, meaning that we only have one asset which depends on the first component of the polynomial process X as it is usually the case in stochastic volatility models. If the Hamburger moment problem for the law of $X_{T,1}$, say μ , admits a unique solution, then the set of all polynomials is dense in $L^2(\mu)$ (see [1, Theorem 2.3.3]) and hence also in $L^1(\mu)$. A sufficient conditions for the uniqueness of a solution to this moment problem is

$$|\mathbb{E}_x[X_{T,1}^k]| \leq C \frac{k!}{R^k}$$

for some constants $C > 0$ and $R > 0$ (see [20, Example X.4]). This condition can be assured by the existence of exponential moments of $X_{T,1}$ around 0, that is, the moment generating function $\mathbb{E}[e^{uX_{T,1}}]$ should be finite for all $u \in (-\varepsilon, \varepsilon)$, which is often satisfied in financial applications.

For illustrative purposes, we have implemented the following affine stochastic volatility model which was initially proposed by Bates [2]. The price process is specified as $S_t = S_0 e^{X_t}$ with dynamics

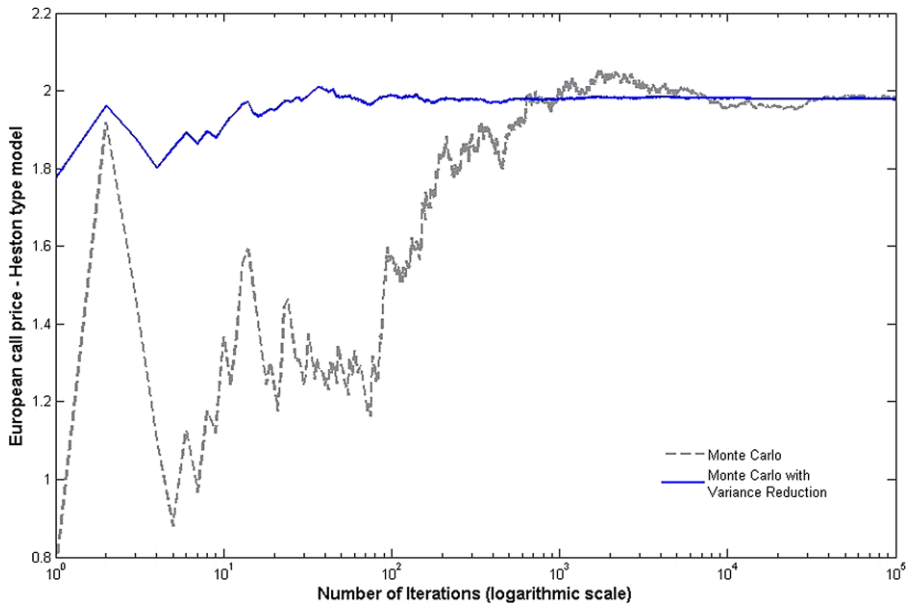
$$\begin{aligned} d \begin{pmatrix} X_t \\ V_t \end{pmatrix} = & \begin{pmatrix} r - \frac{V_t}{2} - \lambda V_t \int_{\mathbb{R}} (e^\xi - 1) F(d\xi) \\ b - \beta V_t \end{pmatrix} dt \\ & + \begin{pmatrix} \sqrt{V_t} & 0 \\ \sigma \rho \sqrt{V_t} & \sigma \sqrt{1 - \rho^2} \sqrt{V_t} \end{pmatrix} \begin{pmatrix} dB_{t,1} \\ dB_{t,2} \end{pmatrix} + \begin{pmatrix} dZ_t \\ 0 \end{pmatrix}, \end{aligned}$$

where B is a 2-dimensional Brownian motion and Z a pure jump process in \mathbb{R} with jump intensity λv and exponentially distributed jump sizes, i.e., $F(\xi) = \frac{1}{c} e^{-\frac{\xi}{c}}$, for some parameter $c \in \mathbb{R}_+$. Figure 1 illustrates the comparison between the Monte Carlo simulation for European call prices with and without variance reduction. In this example we use the parameters as given in Table 1.

The polynomial which we take to approximate the payoff function is of degree 10 and is chosen such as to minimize the approximation error in a certain interval

Table 1 Model parameters

S_0	V_0	Strike	r	b	β	σ	ρ	λ	c
10	0.1	9	0.04	0.08	0.7	0.03	0	1.5	0.05

**Fig. 1** Comparison: Monte Carlo simulation for European call prices with and without variance reduction

(depending on the support of the probability distribution). Concerning computation time, we remark that beside the one-time calculation of the matrix exponential, the only additional computational effort resulting from the use of the control variates is the evaluation of a polynomial in each loop. In our MATLAB code, this causes an increase of computation time of less than 50 % per replication. Observing that one can achieve the same accuracy by using 100 times fewer replications through the polynomial control variates, the computation time (in our MATLAB implementation) can be decreased by a factor of more than 65.

Let us finally remark that our variance reduction technique is of particular interest for affine models for which the generalized Riccati ODEs (see [7]) determining the characteristic function cannot be explicitly solved. Moreover, it can also be applied to derivatives involving several assets, provided that their dynamics are described by a polynomial process. This can simply be done by approximating European payoff functions depending on several variables with multivariate polynomials.

Acknowledgements All authors gratefully acknowledge the support from the FWF grant Y 328 (START prize from the Austrian Science Fund) and from ETH Foundation. Furthermore the authors are grateful for many comments and valuable suggestions by Chris Rogers, which improved our paper a lot.

References

1. Akhiezer, N.I.: The Classical Moment Problem and Some Related Questions in Analysis. Hafner Publishing Co., New York (1965). Translated by N. Kemmer
2. Bates, D.S.: Post-'87 crash fears in the S&P 500 futures option market. *J. Econom.* **94**, 181–238 (2000)
3. Carr, P., Madan, D.: Option valuation using the fast Fourier transform. *J. Comput. Finance* **2**, 61–73 (1998)
4. Chen, L., Filipović, D., Poor, H.V.: Quadratic term structure models for risk-free and defaultable rates. *Math. Finance* **14**, 515–536 (2004)
5. Cheridito, P., Filipović, D., Yor, M.: Equivalent and absolutely continuous measure changes for jump-diffusion processes. *Ann. Appl. Probab.* **15**, 1713–1732 (2005)
6. Çinlar, E., Jacod, J., Protter, P., Sharpe, M.J.: Semimartingales and Markov processes. *Z. Wahrscheinlichkeitstheor. Verw. Geb.* **54**, 161–219 (1980)
7. Duffie, D., Filipović, D., Schachermayer, W.: Affine processes and applications in finance. *Ann. Appl. Probab.* **13**, 984–1053 (2003)
8. Dunkl, C.F.: Hankel transforms associated to finite reflection groups. In: Beals, R.W. (ed.) *Hypergeometric Functions on Domains of Positivity, Jack Polynomials, and Applications*, Tampa, FL, 1991. *Contemp. Math.*, vol. 138, pp. 123–138. Amer. Math. Soc., Providence (1992)
9. Engel, K.-J., Nagel, R.: *One-Parameter Semigroups for Linear Evolution Equations*. Graduate Texts in Mathematics, vol. 194. Springer, New York (2000). With contributions by S. Brendle, M. Campiti, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, S. Romanelli, R. Schnaubelt
10. Forman, J.L., Sørensen, M.: The Pearson diffusions: A class of statistically tractable diffusion processes. *Scand. J. Stat.* **35**, 438–465 (2008)
11. Gallardo, L., Yor, M.: Some remarkable properties of the Dunkl martingales. In: Emery, M., Yor, M. (eds.) *In memoriam Paul-André Meyer: Séminaire de Probabilités XXXIX*. Lecture Notes in Math., vol. 1874, pp. 337–356. Springer, Berlin (2006)
12. Gatheral, J.: *The Volatility Surface: A Practitioner's Guide*. Wiley Finance (2006)
13. Golub, G.H., Van Loan, C.F.: *Matrix Computations*. Johns Hopkins Studies in the Mathematical Sciences, 3rd edn. Johns Hopkins University Press, Baltimore (1996)
14. Gourieroux, C., Jasiak, J.: Multivariate Jacobi process with application to smooth transitions. *J. Econom.* **131**, 475–505 (2006)
15. Jacod, J., Shiryaev, A.N.: *Limit Theorems for Stochastic Processes*. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 288. Springer, Berlin (1987)
16. Jacod, J., Kurtz, T.G., Méléard, S., Protter, P.: The approximate Euler method for Lévy driven stochastic differential equations. *Ann. Inst. Henri Poincaré Probab. Stat.* **41**, 523–558 (2005)
17. Keller-Ressel, M., Schachermayer, W., Teichmann, J.: Affine processes are regular. *Probab. Theory Relat. Fields* **151**, 591–611 (2011)
18. Kunita, H.: Stochastic differential equations based on Lévy processes and stochastic flows of diffeomorphisms. In: *Real and Stochastic Analysis*. Trends Math., pp. 305–373. Birkhäuser, Boston (2004)
19. Moler, C., Van Loan, C.: Nineteen dubious ways to compute the exponential of a matrix. *SIAM Rev.* **20**, 801–836 (1978)
20. Reed, M., Simon, B.: *Methods of Modern Mathematical Physics. II. Fourier Analysis, Self-Adjointness*. Academic Press [Harcourt Brace Jovanovich Publishers], New York (1975)
21. Reimer, M.: *Multivariate Polynomial Approximation*. International Series of Numerical Mathematics, vol. 144. Birkhäuser, Basel (2003)
22. Rogers, L.C.G., Williams, D.: *Diffusions, Markov Processes, and Martingales. Vol. 1 Foundations*. Cambridge Mathematical Library. Cambridge University Press, Cambridge (2000). Reprint of the second (1994) edition
23. Zhou, H.: Itô conditional moment generator and the estimation of short rate processes. *J. Financ. Econom.* **1**, 250–271 (2003)